

# データサイエンスに基づく 医療革命

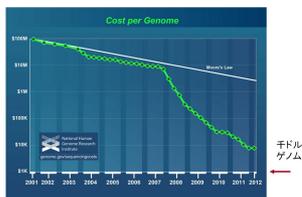
東京医科歯科大学 医療データ科学推進室

田中 博



# 医療ビッグデータ時代の到来

- (1)次世代シーケンサ (Clinical Sequencing)による「ゲノム/オミックス医療」における網羅的分子情報収集/蓄積
- (2) **Biobank/ゲノムコホート普及**による分子・環境情報の蓄積
- (3)**モバイルヘルス(mHealth)** によるWearable センサの連続計測による生理データの蓄積 (unobstructed monitoring)
- (4)医療情報の電子化による**リアルワールドデータ**の大量集積



DNA Sequencing Cost: the National Human Genome Research Institute

急激な大量データの出現  
コストレス化かつ高精度化



ゲノム : 13年→1日(1/5000) 3500億→10万円(1/350万)

医療・ヘルスケアの土台を揺るがす  
ビッグデータ医療時代の到来

# 次世代シーケンサのインパクト

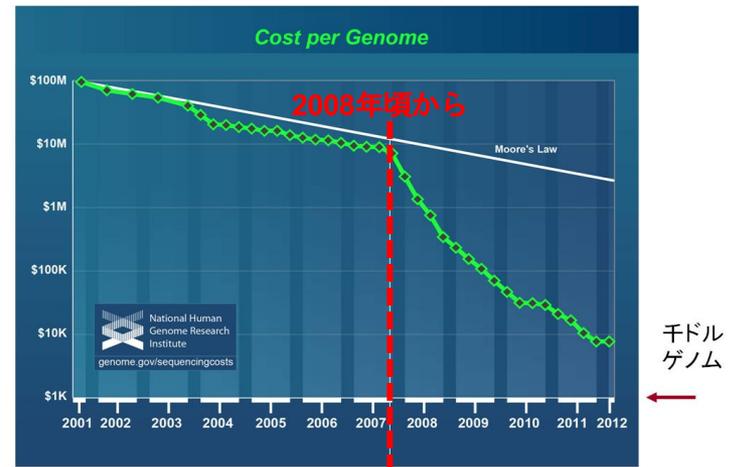
次世代シーケンサを始めとするhigh-throughput分子情報収集の急激な発展

急速な高速化と廉価化 ヒトゲノム解読計画13年,3500億円⇒1日,10万円

2005~ NGS 454 (LS,Roche)  
2007/8~454, Solexa (Illumina),  
SOLiD (LT,TF)  
シーケンス革命



|                  | HiSeq2500       | Ion Proton    |
|------------------|-----------------|---------------|
| 本体価格             | 約1億円            | 約3500万円       |
| モード / チップ        | ハイアウトプット ラピッドラン | Ion Proton I  |
| 解析時間             | 11日             | 27時間          |
| リード長 (bp)        | 2 x 100         | 2 x 150       |
| データ産出量 (Gb)      | 約600            | 約120          |
| 試薬コスト (ヒト1人全ゲノム) | 数十万円            | 不可<br>エクソームのみ |



- Whole genome seq. Whole exsome seq.  
Whole RNA seq. ゲノムワイド性あり  
廉価で高速に収集可能

DNA Sequencing Cost: the National Human Genome Research Institute

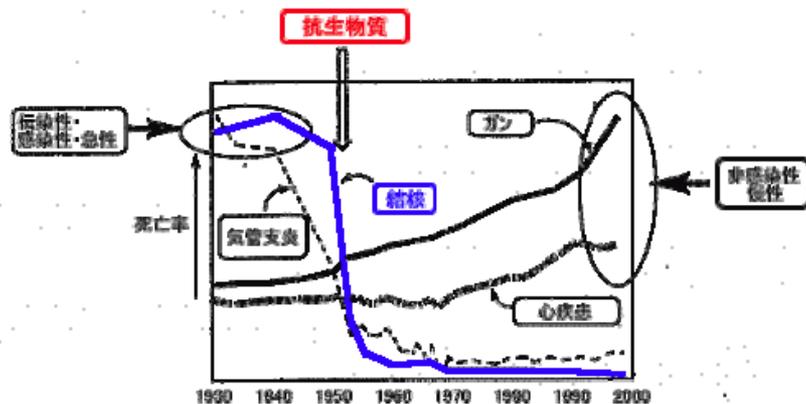
シーケンス革命 2007/8

ゲノム(配列決定)機器の進歩は、計算機のムーアの法則を越えている!

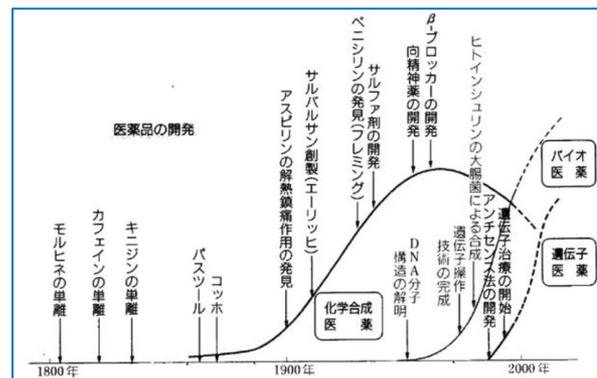
# 根本的革命的性

- 医療における〈ビッグデータ・AI〉は、ここ2～3年の**一時的ブーム**ではない。
  - 医学研究の進め方、医療の向上に関する枠組みについての**何十年と続く根底的な革命**の開始。
- ビッグデータ・AIによる**医療の第3次革命**
    - 抗生物質の発明による細菌性感染症の克服（**第1革命**）、分子生命科学による分子標的薬・抗体医薬などの発明（**第2革命**）に次ぐ、**第3次革命**。

第1世代(1930～1970)



第3世代 (2010～2040)

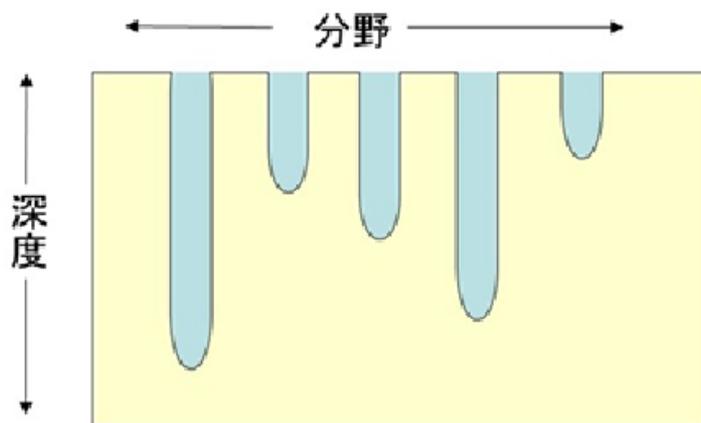


# 第3次革命の基軸

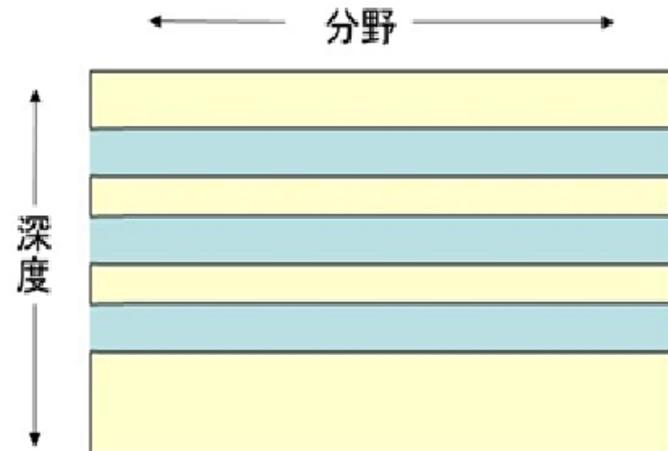
## データ駆動型医学(Data-driven Medicine)

- 従来の医学は医学者が研究関心や社会的要請に従って従前の知識を前提として、新たな仮説を立て、実験によって検証していく
- これまでの医学：仮説駆動的医学 (hypothesis-driven medicine)
  - 真理性は実験の仮説反駁可能性 (Popper)
- ビッグデータ：全体表出性；網羅的データからAIで知識を発見できる
- 生命情報データのゲノムワイド性(GWAS, Omics, 次世代シーケンサ)
- Churing-Nobel Challenge (AIがノーベル賞を取る)

従前研究を前提、研究関心、医療の社会的要請、孤発的医学知識  
穴だらけ、研究深度もまちまち



網羅的データに準拠した医学研究  
網羅性・分野横断的知識  
同一の深度を持つ知識



# ビッグデータ医療はパラダイム 変換を実現可能にする

ビッグデータの医療・創薬への効果(20~30年)

I ゲノム・オミックス情報

疾患成立機序の解明

II 大規模バイオバンク情報

長期生涯疾患過程の解明

III モバイルヘルス情報

短期生涯疾患過程の解明

IV リアルワールド情報

生涯の健康/疾患データ

Disease Big Data



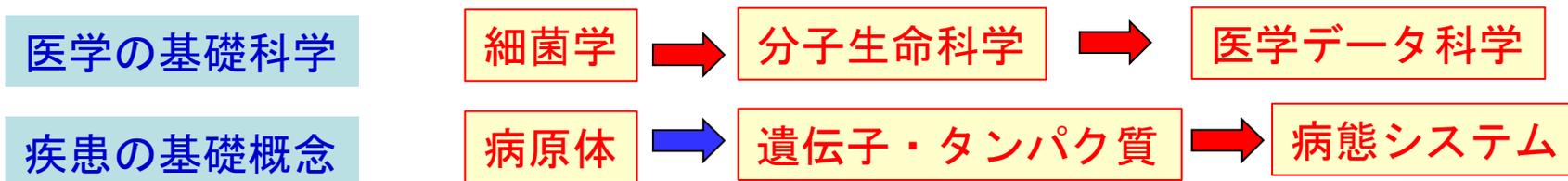
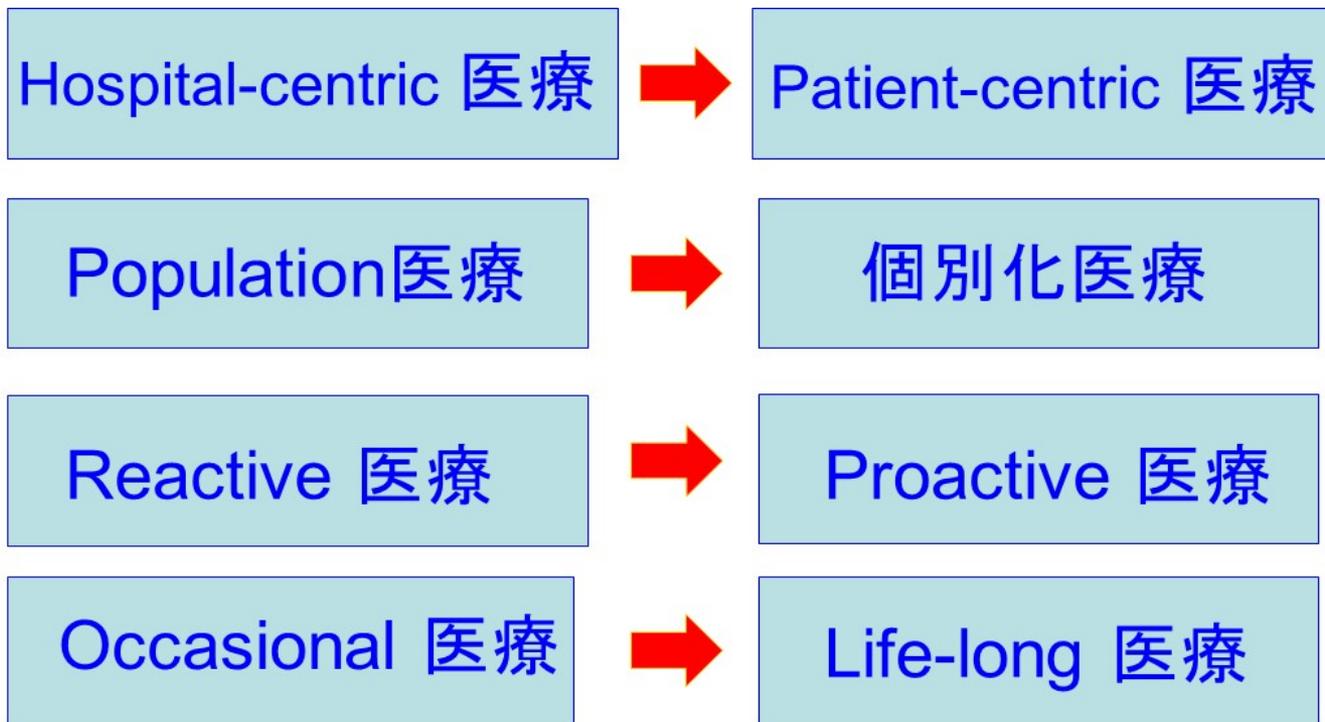
個別化医療・先制医療

生涯医療・予知医療・参加医療



Life-long Big Data

# 医療の長期的パラダイム変換



## ■ 医療・創薬データサイエンスコンソーシアムの運営体制

### 連携機関(企業)

- ◆プログラム受講者(企業人材コース)の参加
- ◆企業インターンシップ(博士人材コース対象)の実施
- ◆大学院生・ポストドクター等のリクルート

製薬、化学、情報等21社の加入(2018年度実績)

## 医療・創薬データサイエンス コンソーシアム

代表機関：東京医科歯科大学

医療・創薬データサイエンスコンソーシアム事務局：  
株式会社シード・プランニング

一般社団法人バイオ産業情報化コンソーシアム

### 高等教育機関

- ◆プログラム受講者(博士人材コース・対象は大学院生・ポストドクター等)の参加
- ◆人材育成プログラムへの講師派遣
- ◆研修(企業人材コース対象)の実施

東北大学大学院薬学研究科、  
東北メディカル・メガバンク機構  
慶應義塾大学大学院薬学研究科、  
政策・メディア研究科、先端生命科学研究所  
東京理科大学大学院薬学研究科

### 公的研究機関

- ◆人材育成プログラムへの講師派遣
- ◆研修(企業人材コース対象)の実施

産業技術総合研究所人工知能研究センター、  
創薬分子プロファイリング研究センター  
国立精神・神経医療研究センター  
国立国際医療研究センター  
公益財団法人がん研究会

### 連携機関

| 平成29年度(22社)<br>年会費50万円 | 平成30年度(21社)<br>年会費200万円  |
|------------------------|--------------------------|
| 小野薬品工業株式会社             | 小野薬品工業株式会社               |
| 中外製薬株式会社               | 中外製薬株式会社                 |
| 旭化成ファーマ株式会社            | 旭化成ファーマ株式会社              |
| アステラス製薬株式会社            | アステラス製薬株式会社              |
| 田辺三菱製薬株式会社             | 田辺三菱製薬株式会社               |
| 第一三共株式会社               | 第一三共株式会社                 |
| デンカ株式会社                | デンカ株式会社                  |
| 帝人ファーマ株式会社             | 帝人ファーマ株式会社               |
| 日本電気株式会社               | 日本電気株式会社                 |
| 富士通株式会社                | 富士通株式会社                  |
| 株式会社シーエーシー             | 株式会社シーエーシー               |
| 日本たばこ産業株式会社            | 日本たばこ産業株式会社              |
| 科研製薬株式会社               | 科研製薬株式会社                 |
| 協和発酵キリン株式会社            | 協和発酵キリン株式会社              |
| ファイザー株式会社              | ファイザー株式会社                |
| 武田薬品工業株式会社             | 日本ケミファ株式会社               |
| 大塚製薬株式会社               | 大日本住友製薬株式会社              |
| 大鵬薬品工業株式会社             | Meiji Seika ファルマ株式<br>会社 |
| 木村情報技術株式会社             | 株式会社DeNAライフサイ<br>エンス     |
| 株式会社ファーストアセ<br>ント      | 興和株式会社                   |
| 味の素株式会社                | 株式会社日立製作所                |
| 株式会社CACクロア             |                          |

- \* 製薬・化学・情報等の民間企業が連携機関として参画
- \* 博士人材を対象とした企業インターンシップの実施。大学院生・ポストドクター等の採用活動も行う



# コアカリキュラム

## ■ 入門編・基礎編を増設、実習時間を拡充

### 2019年度 カリキュラム概要



大阪大学  
連携分

大阪大学、早稲田  
大学とカリキュラム  
構築の点で連携

早稲田大学  
連携分

#### 入門編 (導入編)

◇講義

- データサイエンス入門
- 統計リテラシーシリーズ  $\alpha$
- 統計リテラシーシリーズ  $\beta$
- 統計リテラシーシリーズ  $\gamma$

※以上は博士人材コース  
限定科目。  
• データサイエンスのための  
基礎数学

◇実習

- Linux 実習

※赤字の講義・実習は  
2019年度に新設

#### 基礎編

◇講義

- データサイエンス基礎I
- データサイエンス基礎II
- マネジメント特論
- 知的財産特論
- 生命倫理

◇実習

- プログラミング実習I (R)
- プログラミング実習II (Python)
- 医療統計解析実習
- SQL入門
- 可視化&GUI実装実習
- NGS解析入門
- 深層学習の基礎技術

#### 応用編

◇講義

- 人工知能・機械学習概論
- ゲノム統計学概論
- 疾患データ科学特論
- AI創薬/現代医薬統計学  
特論他

◇実習

- オミックス解析
- テキストマイニング
- 統計モデリング
- ネットワーク解析
- オープンデータ入門
- IoT・時系列解析
- ブロックチェーン入門
- GWAS 解析
- AI 創薬・ケイノマティクス入門

#### 実践編

◇研修 (企業人材コース対象)

◇企業インターンシップ  
(博士人材コース対象)

◇ワークショップ (共通)

見習いレベル

一人前レベル

棟梁レベル

# 医療データサイエンスを 東京医科歯科大学で学び

1. 医療データ科学推進室（田中博室長）
2. M&Dデータ科学センター（宮野悟センター長）

## 産学連携を実現しよう

ゲノム・オミックス、モバイルヘルス、  
リアルワールドデータのAI利用  
すでに1社とは共同研究中